Density estimates and concentration inequalities with Malliavin calculus

Ivan Nourdin, Frederi Viens

To cite this version:

Ivan Nourdin, Frederi Viens. Density estimates and concentration inequalities with Malliavin calculus. 2008. hal-00311755

HAL Id: hal-00311755

https://hal.science/hal-00311755

Preprint submitted on 20 Aug 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Density estimates and concentration inequalities with Malliavin calculus

Ivan Nourdin[*] and Frederi G. Viens[†]
Université Paris 6 and *Purdue University*

Abstract

We show how to use the Malliavin calculus to obtain density estimates of the law of general centered random variables. In particular, under a non-degeneracy condition, we prove and use a new formula for the density ρ of a random variable Z which is measurable and differentiable with respect to a given isonormal Gaussian process. Among other results, we apply our techniques to bound the density of the maximum of a general Gaussian process from above and below; several new results ensue, including improvements on the so-called Borell-Sudakov inequality. We then explain what can be done when one is only interested in or capable of deriving concentration inequalities, i.e. tail bounds from above or below but not necessarily both simultaneously.

Key words: Malliavin calculus; density estimates; concentration inequalities; fractional Brownian motion; Borell-Sudakov inequality; suprema of Gaussian processes.

2000 Mathematics Subject Classification: 60G15; 60H07.

1 Introduction

Let N be a zero-mean Gaussian random vector, with covariance matrix $K \in \mathscr{S}_n^+(\mathbb{R})$. Set $\sigma_{\max}^2 := \max_i K_{ii}$, and consider

$$Z = \max_{1\leqslant i\leqslant n} N_i - E\big(\max_{1\leqslant i\leqslant n} N_i\big). \tag{1.1}$$

[*]Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie, Boîte courrier 188, 4 Place Jussieu, 75252 Paris Cedex 5, France, ivan.nourdin@upmc.fr

[†]Dept. Statistics and Dept. Mathematics, Purdue University, 150 N. University St., West Lafayette, IN 47907-2067, USA, viens@purdue.edu

It is well-known, see e.g. Vitale [16], that for all $z > 0$,

$$P\big(Z \geqslant z\big) \leqslant \exp\left(-\frac{z^2}{2\,\sigma_{\max}^2}\right) \quad \text{if } z > 0. \tag{1.2}$$

The corresponding left-tail probability bound analogue of (1.2) also holds, see e.g. Borell [2]:

$$P\big(Z \leqslant -z\big) \leqslant \exp\left(-\frac{z^2}{2\,\sigma_{\max}^2}\right) \quad \text{if } z < 0. \tag{1.3}$$

Of course, we can combine (1.2) and (1.3) to get,

$$P\big(|Z| \geqslant z\big) \leqslant 2\exp\left(-\frac{z^2}{2\,\sigma_{\max}^2}\right) \quad \text{if } z > 0. \tag{1.4}$$

Inequality (1.4) is a special case of bounds for more general Gaussian fields. Such bounds are often collectively known as Borell-Sudakov inequalities. These can be extended much beyond the Gaussian realm; see for instance the book of Ledoux and Talagrand [10]. Yet these Borell-Sudakov inequalities can still be improved, even in the Gaussian framework; this is one of the things we will illustrate in this paper.

Inequality (1.4) is also a special case of results based on almost sure bounds on a random field's Malliavin derivatives, see Viens and Vizcarra [15]. While that paper uncovered a new way to relate scales of regularity and fractional exponential moment conditions with iterated Malliavin derivatives, it failed to realize how best to use these derivatives when seeking basic estimates such as (1.4). In the present paper, our aim is to explain how to use Malliavin calculus more efficiently than in [15] in order to obtain bounds like (1.2) or (1.3), and even often much better. For instance, by applying our machinery to Z defined by (1.1), we obtain the following.

Proposition 1.1 *With N and Z as above, if $\sigma_{\min}^2 := \min_{i,j} K_{ij} > 0$, with $\sigma_{\max}^2 := \max_i K_{ii}$, the density ρ of Z exists and satisfies, for almost all $z \in \mathbb{R}$,*

$$\frac{E|Z|}{2\sigma_{\max}^2}\exp\left(-\frac{z^2}{2\,\sigma_{\min}^2}\right) \leqslant \rho(z) \leqslant \frac{E|Z|}{2\sigma_{\min}^2}\exp\left(-\frac{z^2}{2\,\sigma_{\max}^2}\right). \tag{1.5}$$

This proposition generalizes immediately (see Proposition 3.11 in Section 3 below) to the case of processes defined on an interval $[a, b] \subset \mathbb{R}$. To our knowledge, that result is the first instance where the density of the maximum of a general Gaussian process is estimated from above and below. As an explicit application, let us mention the following result, concerning the centered maximum of a fractional Brownian motion (fBm), which is proved at the end of Section 3.

Proposition 1.2 *Let $b > a > 0$, and $B = (B_t, t \geqslant 0)$ be a fractional Brownian motion with Hurst index $H \in (1/2, 1)$. Then the random variable $Z = \sup_{[a,b]} B - E\big(\sup_{[a,b]} B\big)$ has a density ρ satisfying, for almost all $z \in \mathbb{R}$:*

$$\frac{E|Z|}{2b^{2H}}\, e^{-\frac{z^2}{2\,a^{2H}}} \leqslant \rho(z) \leqslant \frac{E|Z|}{2a^{2H}}\, e^{-\frac{z^2}{2\,b^{2H}}}. \tag{1.6}$$

Of course, the interest of this result lies in the fact that the exact distribution of $\sup_{[a,b]} B$ is still an open problem when $H \neq 1/2$. Moreover, note that introducing a degeneracy in the covariances for stochastic processes such as fBm has dire consequences on their supremas' tails; for instance, with $a = 0$, Z has no left hand tail, since $Z \geqslant -E\big(\sup_{[0,b]} B\big)$ a.s., and therefore ρ is zero for z small enough.

Density estimates of the type (1.5) may be used immediately to derive tail estimates by combining simple integration with the following classical inequalities:

$$\frac{z}{1+z^2}\, e^{-\frac{z^2}{2}} \leqslant \int_z^\infty e^{-\frac{y^2}{2}} dy \leqslant \frac{1}{z}\, e^{-\frac{z^2}{2}} \quad \text{for all } z > 0.$$

The two tails of the supremum of a Gaussian vector or process are typically not symmetric, and neither are the methods for estimating them; this poses a problem for the techniques used in [16] and [2], and for ours. Let us therefore first derive some results by hand. For a lower bound on the right-hand tail of Z, no heavy machinery is necessary. Indeed let $i_0 = \arg\max_i K_{ii}$ and $\mu = E\big(\max N_i\big) > 0$. Then, for $z > 0$,

$$P\big(Z \geqslant z\big) \geqslant P\big(N_{i_0} \geqslant \mu + z\big) \geqslant \frac{1}{\sqrt{2\pi}}\, \frac{(\mu+z)^2}{\sigma_{\max}^2 + (\mu+z)^2}\, e^{-\frac{(\mu+z)^2}{2\,\sigma_{\max}^2}}. \tag{1.7}$$

A nearly identical argument leads to the following upper bound on the left-hand tail of Z: for $z > 0$,

$$P(Z \leqslant -z) \leqslant \frac{\min_i \sqrt{K_{ii}}}{\sqrt{2\pi}(z-\mu)} \exp\left(-\frac{(z-\mu)^2}{2\min_i K_{ii}}\right). \tag{1.8}$$

This improves Borell's inequality (1.3) asymptotically.

By using the techniques in our article, the density estimates in (1.5) allow us to obtain a new lower bound result on Z's left hand tail, and to improve the classical right-hand tail result of (1.2). We have for the right-hand tail

$$\frac{E|Z|\,\sigma_{\min}^2}{2\,\sigma_{\max}^2}\, \frac{z}{\sigma_{\min}^2 + z^2} \exp\left(-\frac{z^2}{2\,\sigma_{\min}^2}\right) \leqslant P\big(Z \geqslant z\big) \leqslant \frac{E|Z|\,\sigma_{\max}^2}{2\,\sigma_{\min}^2}\, \frac{1}{z} \exp\left(-\frac{z^2}{2\,\sigma_{\max}^2}\right) \tag{1.9}$$

if $z > 0$, and one notes that the above right-hand side goes (slightly) faster to zero than (1.4), because of the presence of the factor z^{-1}; yet the lower bound is less sharp than (1.7) for large z. The first and last expressions in (1.9) are also lower and upper bounds for the left-hand tail $P\big(Z \leqslant -z\big)$. To the best of our knowledge, the lower bound is new; the upper bound is less sharp than (1.8) for large z.

Let us now cite some works which are related to ours, insofar as some of the preoccupations and techniques are similar. In [6], Houdré and Privault prove concentration inequalities for functionals of Wiener and Poisson spaces: they have discovered almost-sure conditions on expressions involving Malliavin derivatives which guarantee upper bounds on the tails of their functionals. This is similar to the upper bound portion of our work in Section 4, and closer yet to the first-chaos portion of the work in [15]; they do not, however, address lower bound issues, nor do they have any claims regarding densities.

Decreusefond and Nualart [5] obtain, by means of the Malliavin calculus, estimates for the Laplace transform of the hitting times of any general Gaussian process; they define a monotonicity condition on the covariance function of such a process under which this Laplace transform is bounded above by that of standard Brownian motion; similarly to how we derive upper tail estimates of Gaussian type from our analysis, they derive the finiteness of some moments by comparison to the Brownian case. However, as in [6], reference [5] does not address issues of densities or of lower bounds.

General lower bound results on densities are few and far between. The case of uniformly elliptic diffusions was treated in a series of papers by Kusuoka and Stroock: see [9]. This was generalized by Kohatsu-Higa [8] in Wiener space via the concept of uniformly elliptic random variables; these random variables proved to be well-adapted to studying diffusion equations. E. Nualart [13] showed that fractional exponential moments for a divergence-integral quantity known to be useful for bounding densities from above (see formula (1.10) below), can also be useful for deriving a scale of exponential lower bounds on densities; the scale includes Gaussian lower bounds. However, in all these works, the applications are largely restricted to diffusions.

We now introduce our general setting which will allow to prove (1.5)-(1.6) and several other results. We consider a centered isonormal Gaussian process $X = \{X(h) : h \in \mathfrak{H}\}$ defined on a real separable Hilbert space $\mathfrak{H}$. This just means that X is a collection of centered and jointly Gaussian random variables indexed by the elements of $\mathfrak{H}$, defined on some probability space $(\Omega, \mathscr{F}, P)$ and such that, for every $h, g \in \mathfrak{H}$,

$$E\big(X(h)X(g)\big) = \langle h, g\rangle_{\mathfrak{H}}.$$

As usual in Malliavin calculus, we use the following notation (see Section 2 for precise definitions):

- $L^2(\Omega, \mathscr{F}, P)$ is the space of square-integrable functionals of X. This means in particular that $\mathscr{F}$ is the σ-field generated by X;

- $\mathbb{D}^{1,2}$ is the domain of the Malliavin derivative operator D with respect to X. Roughly speaking, it is the subset of random variables in $L^2(\Omega, \mathscr{F}, P)$ whose Malliavin derivative is also in $L^2(\Omega, \mathscr{F}, P)$;

- Domδ is the domain of the divergence operator δ. This operator will really only play a marginal role in our study; it is simply used in order to simplify some proof arguments, and for comparison purposes.

From now on, Z will always denote a random variable of $\mathbb{D}^{1,2}$ with *zero mean*. Recall that its derivative DZ is a random element with values in $\mathfrak{H}$. The following result on the density of a random variable is a well-known fact of the Malliavin calculus: if $DZ/\|DZ\|_{\mathfrak{H}}^2$ belongs to Domδ, then Z has a continuous and bounded density ρ given, for all $z \in \mathbb{R}$, by

$$\rho(z) = E\left[\mathbf{1}_{(z,+\infty]}(Z)\,\delta\left(\frac{DZ}{\|DZ\|_{\mathfrak{H}}^2}\right)\right]. \tag{1.10}$$

From this expression, it is sometimes possible to deduce *upper* bounds for ρ. Several examples are detailed in Section 2.1.1 of Nualart's book [12]. Note the following two points, however: (a) it is not clear whether it is at all possible to prove (1.5) by using (1.10); (b) more generally it appears to be just as difficult to deduce any *lower*-bound relations on the density ρ of any random variable via (1.10).

Herein we prove a new general formula for ρ, from which we easily deduce (1.5) for instance. For Z a mean-zero r.v. in $\mathbb{D}^{1,2}$, define the function $g : \mathbb{R} \to \mathbb{R}$ almost everywhere by

$$g(z) = g_Z(z) := E\left(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}\big|\, Z = z\right). \tag{1.11}$$

The L appearing here is the so-called generator of the Ornstein-Uhlenbeck semigroup, defined in the next section. We drop the subscript $_Z$ from g_Z in this article, since each example herein refers to only one r.v. Z at a time. By [11, Proposition 3.9], g is non-negative on the support of Z. Under some general conditions on Z (see Theorem 3.1 for a precise statement), the density ρ of Z is given by the following new formula, for any z in Z's support:

$$P\left(Z \in dz\right) = \rho(z)dz = \frac{E|Z|}{2g(z)}\exp\left(-\int_0^z \frac{x\,dx}{g(x)}\right)dz. \tag{1.12}$$

The key point in our approach is that it is possible, in many cases, to estimate the quantity $g(z)$ in (1.11) rather precisely. In particular, we will make systematic use of the following consequence of the Mehler formula (see Remark 3.6 in [11]), also proved herein (Proposition 3.5):

$$g(z) = \int_0^\infty e^{-u}\,\mathbf{E}\big(\langle \Phi_Z(X), \Phi_Z(e^{-u}X + \sqrt{1-e^{-2u}}X')\rangle_{\mathfrak{H}} | Z = z\big)du.$$

In this formula, the mapping $\Phi_Z : \mathbb{R}^{\mathfrak{H}} \to \mathfrak{H}$ is defined $P \circ X^{-1}$-almost surely through the identity $DZ = \Phi_Z(X)$, while X', which stands for an independent copy of X, is such that X and X' are defined on the product probability space $(\Omega\times\Omega', \mathscr{F}\otimes\mathscr{F}', P\times P')$; $\mathbf{E}$ denotes the mathematical expectation with respect to $P \times P'$. This formula for g then allows, in many cases, to obtain via (1.12) a lower and an upper bound on ρ simultaneously. We refer the reader to Corollary 3.6 and the examples in Section 3, and in particular to the second and fourth examples, which are the proofs of Proposition 1.1 and Proposition 1.2 respectively. At this stage, let us note however that it is not possible to obtain only a lower bound, or only an upper bound, using formula (1.12). Indeed, one can see that one needs to control g simultaneously from above and below to get the technique to work.

In the second main part of the paper (Section 4), we explain what can be done when one only knows how to bound g from one direction or the other, but not both simultaneously. Note that one is precisely in this situation when one seeks to prove the inequalities (1.2) and (1.3). These will be a simple consequence of a more general upper bound result (Theorem 4.1) in Section 4.

As another application of Theorem 4.1, the following result concerns a functional of fractional Brownian motion.

Proposition 1.3 *Let $B = \{B_t,\, t \in [0,T]\}$ be a fractional Brownian motion with Hurst index $H \in (0,1)$. Then, denoting $c_H = H + 1/2$, we have, for any $z > 0$:*

$$P\left(\int_0^T B_u^2 du \geqslant z + T^{2H+1}/(2c_H)\right) \leqslant \exp\left(-\frac{c_H^2\ z^2}{2c_H T^{2H+1} z + T^{4H+2}}\right).$$

Of course, the interest of this result lies in the fact that the exact distribution of $\int_0^T B_u^2 du$ is still an open problem when $H \neq 1/2$. With respect to the classical result by Borell [1] (which would give a bound like $\exp(-Cz)$), observe here that, as in Chatterjee [3], we get a kind of "continuous" transition from Gaussian to exponential tails. The behavior for large z is always of exponential type. At the end of this article, we take up the issue of finding a lower bound which might be commensurate with the upper bound above; our Malliavin calculus techniques fail here, but we are still able to derive an interesting result by hand, see (4.28).

Section 4 also contains a lower bound result, Theorem 4.2, again based on the quantity $\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}$ via the function g in (1.11). This quantity was introduced recently in [11] for the purpose of using Stein's method in order to show that the standard deviation of $\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}$ provides an error bound of the normal approximation of Z, see also Remark 3.2 below. Here, in Theorem 4.2 and in Theorem 4.1 as a special case ($\alpha = 0$ therein), $g(Z) = E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z)$ can be instead assumed to be bounded either above *or* below *almost surely* by a constant; this constant's role is to be a measure of the variance of Z, and more specifically to ensure that the tail of Z is bounded either above or below by a normal tail with that constant as its variance. Our Section 4 can thus be thought as a way to extend the phenomena described in [11] when comparison with the normal distribution can only be expected to go one way. Theorem 4.2 shows that we may have no control over how heavy the tail of Z may be (beyond the existence of a second moment), but the condition $g(Z) \geqslant \sigma^2 > 0$ essentially guarantees that it has to be no less heavy than a Gaussian tail with variance σ^2.

We finish this description of our results by stressing again that, whether in Sections 3 or 4, we present many examples where the quantities $\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}$ and $\langle \Phi_Z(X), \Phi_Z(e^{-u}X+$

$\sqrt{1-e^{-2u}}X'))\rangle_{\mathfrak{H}}$ are computed and estimated easily, by hand and/or via Proposition 3.5. The advantage over formulas such as (1.10), which involve the unwieldy divergence operator δ, should be clear.

The rest of the paper is organized as follows. In Section 2, we recall the notions of Malliavin calculus that we need in order to perform our proofs. In Section 3, we state and discuss our density estimates. Section 4 deals with concentration inequalities, i.e. tail estimates.

2 Some elements of Malliavin calculus

We follow Nualart's book [12]. As stated in the introduction, we denote by X a centered isonormal Gaussian process over a real separable Hilbert space $\mathfrak{H}$. Let $\mathscr{F}$ be the σ-field generated by X. It is well-known that any random variable Z belonging to $L^2(\Omega,\mathscr{F},P)$ admits the following chaos expansion:

$$Z=\sum_{m=0}^{\infty} I_m(f_m), \tag{2.13}$$

where $I_0(f_0)=E(Z)$, the series converges in $L^2(\Omega)$ and the kernels $f_m\in\mathfrak{H}^{\odot m}$, $m\geqslant 1$, are uniquely determined by Z. In the particular case where $\mathfrak{H}=L^2(A,\mathscr{A},\mu)$, for $(A,\mathscr{A})$ a measurable space and μ a σ-finite and non-atomic measure, one has that $\mathfrak{H}^{\odot m}=L^2_s(A^m,\mathscr{A}^{\otimes m},\mu^{\otimes m})$ is the space of symmetric and square integrable functions on A^m and, for every $f\in\mathfrak{H}^{\odot m}$, $I_m(f)$ coincides with the multiple Wiener-Itô integral of order m of f with respect to X. For every $m\geqslant 0$, we write J_m to indicate the orthogonal projection operator on the mth Wiener chaos associated with X. That is, if $Z\in L^2(\Omega,\mathscr{F},P)$ is as in (2.13), then $J_mF=I_m(f_m)$ for every $m\geqslant 0$.

Let $\mathscr{S}$ be the set of all smooth cylindrical random variables of the form

$$Z=g\big(X(\phi_1),\ldots,X(\phi_n)\big)$$

where $n\geqslant 1$, $g:\mathbb{R}^n\to\mathbb{R}$ is a smooth function with compact support and $\phi_i\in\mathfrak{H}$. The Malliavin derivative of Z with respect to X is the element of $L^2(\Omega,\mathfrak{H})$ defined as

$$DZ=\sum_{i=1}^{n}\frac{\partial g}{\partial x_i}\big(X(\phi_1),\ldots,X(\phi_n)\big)\phi_i.$$

In particular, $DX(h)=h$ for every $h\in\mathfrak{H}$. By iteration, one can define the mth derivative D^mZ (which is an element of $L^2(\Omega,\mathfrak{H}^{\odot m})$) for every $m\geqslant 2$. As usual, for $m\geqslant 1$, $\mathbb{D}^{m,2}$ denotes the closure of $\mathscr{S}$ with respect to the norm $\|\cdot\|_{m,2}$, defined by the relation

$$\|Z\|^2_{m,2}=E(Z^2)+\sum_{i=1}^{m}E\big(\|D^iZ\|^2_{\mathfrak{H}^{\otimes i}}\big).$$

Note that a random variable Z as in (2.13) is in $\mathbb{D}^{1,2}$ if and only if

$$\sum_{m=1}^{\infty} m\, m!\, \|f_m\|^2_{\mathfrak{H}^{\otimes m}} < \infty,$$

and, in this case, $E\big(\|DZ\|^2_{\mathfrak{H}}\big) = \sum_{m\geqslant 1} m\, m!\, \|f_m\|^2_{\mathfrak{H}^{\otimes m}}$. If $\mathfrak{H} = L^2(A, \mathscr{A}, \mu)$ (with μ non-atomic), then the derivative of a random variable Z as in (2.13) can be identified with the element of $L^2(A \times \Omega)$ given by

$$D_a Z = \sum_{m=1}^{\infty} m I_{m-1}\big(f_m(\cdot, a)\big), \quad a \in A.$$

The Malliavin derivative D satisfies the following chain rule. If $\varphi : \mathbb{R}^n \to \mathbb{R}$ is of class $\mathscr{C}^1$ with bounded derivatives, and if $\{Z_i\}_{i=1,\ldots,n}$ is a vector of elements of $\mathbb{D}^{1,2}$, then $\varphi(Z_1, \ldots, Z_n) \in \mathbb{D}^{1,2}$ and

$$D\,\varphi(Z_1, \ldots, Z_n) = \sum_{i=1}^{n} \frac{\partial \varphi}{\partial x_i}(Z_1, \ldots, Z_n) DZ_i. \tag{2.14}$$

Formula (2.14) still holds when φ is only Lipshitz but the law of $(Z_1, \ldots, Z_n)$ has a density with respect to the Lebesgue measure on $\mathbb{R}^n$ (see e.g. Proposition 1.2.3 in [12]).

We denote by δ the adjoint of the operator D, also called the divergence operator. A random element $u \in L^2(\Omega, \mathfrak{H})$ belongs to the domain of δ, denoted by Domδ, if and only if it satisfies

$$\big|E\langle DZ, u\rangle_{\mathfrak{H}}\big| \leqslant c_u\, E(Z^2)^{1/2} \quad \text{for any } Z \in \mathscr{S},$$

where c_u is a constant depending only on u. If $u \in$ Domδ, then the random variable $\delta(u)$ is uniquely defined by the duality relationship

$$E(Z\delta(u)) = E\langle DZ, u\rangle_{\mathfrak{H}}, \tag{2.15}$$

which holds for every $Z \in \mathbb{D}^{1,2}$.

The operator L is defined through the projection operators as $L = \sum_{m=0}^{\infty} -mJ_m$, and is called the generator of the Ornstein-Uhlenbeck semigroup. It satisfies the following crucial property. A random variable Z is an element of Dom$L$ ($= \mathbb{D}^{2,2}$) if and only if $Z \in$ DomδD (i.e. $Z \in \mathbb{D}^{1,2}$ and $DZ \in$ Domδ), and in this case:

$$\delta DZ = -LZ. \tag{2.16}$$

We also define the operator L^{-1}, which is the inverse of L, as follows. For every $Z \in L^2(\Omega, \mathscr{F}, P)$, we set $L^{-1}Z = \sum_{m\geqslant 1} -\frac{1}{m} J_m(Z)$. Note that L^{-1} is an operator with values in $\mathbb{D}^{2,2}$, and that $LL^{-1}Z = Z - E(Z)$ for any $Z \in L^2(\Omega, \mathscr{F}, P)$, so that L^{-1} does act as L's inverse for centered r.v.'s.

The family $(T_u,\, u \geqslant 0)$ of operators is defined as $T_u = \sum_{m=0}^{\infty} e^{-mu} J_m$, and is called the Orstein-Uhlenbeck semigroup. Assume that the process X', which stands for an independent copy of X, is such that X and X' are defined on the product probability space $(\Omega\times\Omega', \mathscr{F}\otimes\mathscr{F}', P\times P')$. Given a random variable $Z\in\mathbb{D}^{1,2}$, we can write $DZ = \Phi_Z(X)$, where Φ_Z is a measurable mapping from $\mathbb{R}^{\mathfrak{H}}$ to $\mathfrak{H}$, determined $P\circ X^{-1}$-almost surely. Then, for any $u \geqslant 0$, we have the so-called Mehler formula:

$$T_u(DZ) = E'\big(\Phi_Z(e^{-u}X + \sqrt{1-e^{-2u}}X')\big), \tag{2.17}$$

where E' denotes the mathematical expectation with respect to the probability P'.

3 Density estimates

For $Z\in\mathbb{D}^{1,2}$ with zero mean, recall the function g introduced in the introduction in (1.11):

$$g(z) = E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z=z).$$

It is useful to keep in mind throughout this paper that, by [11, Proposition 3.9], $g(z)\geqslant 0$ on the support of Z. In this section, we further assume that g is bounded away from 0.

3.1 General formulae and estimates

We begin with the following theorem, which will be key in the sequel.

Theorem 3.1 *Let $Z\in\mathbb{D}^{1,2}$ with zero mean, and g as above. Assume that there exists $\sigma_{\min}>0$ such that*

$$g(Z)\geqslant \sigma_{\min}^2 \quad \textit{almost surely.} \tag{3.18}$$

Then Z has a density ρ, its support is $\mathbb{R}$ and we have, almost everywhere:

$$\rho(z) = \frac{E\,|Z|}{2g(z)}\exp\left(-\int_0^z \frac{x\,dx}{g(x)}\right). \tag{3.19}$$

Proof. We split the proof into several steps.

Step 1: An integration by parts formula. For any $f:\mathbb{R}\to\mathbb{R}$ of class $\mathscr{C}^1$ with bounded derivative, we have

$$\begin{aligned} E\big(Zf(Z)\big) &= E\big(LL^{-1}Zf(Z)\big) = E\big(\delta D(-L^{-1}Z)f(Z)\big) \quad \text{by (2.16)}\\ &= E\big(\langle Df(Z), -DL^{-1}Z\rangle_{\mathfrak{H}}\big) \quad \text{by (2.15)}\\ &= E\big(f'(Z)\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}\big) \quad \text{by (2.14)}. \end{aligned} \tag{3.20}$$

Step 2: Existence of the density. Fix $a < b$ in $\mathbb{R}$. For any $\varepsilon > 0$, consider a $\mathscr{C}^\infty$-function $\varphi_\varepsilon : \mathbb{R} \to [0,1]$ such that $\varphi_\varepsilon(z) = 1$ if $z \in [a,b]$ and $\varphi_\varepsilon(z) = 0$ if $z < a - \varepsilon$ or $z > b + \varepsilon$. We set $\psi_\varepsilon(z) = \int_{-\infty}^{z} \varphi_\varepsilon(y)dy$ for any $z \in \mathbb{R}$. Then, we can write

$$\begin{aligned}
P(a \leqslant Z \leqslant b) &= E\big(\mathbf{1}_{[a,b]}(Z)\big) \\
&\leqslant \sigma_{\min}^{-2}\, E\big(\mathbf{1}_{[a,b]}(Z) E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}} | Z)\big) \quad \text{by assumption (3.18)} \\
&= \sigma_{\min}^{-2}\, E\big(\mathbf{1}_{[a,b]}(Z) \langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}\big) \\
&= \sigma_{\min}^{-2}\, E\big(\liminf_{\varepsilon \to 0} \varphi_\varepsilon(Z) \langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}\big) \\
&\leqslant \sigma_{\min}^{-2}\, \liminf_{\varepsilon \to 0} E\big(\varphi_\varepsilon(Z) \langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}\big) \quad \text{by Fatou's inequality} \\
&= \sigma_{\min}^{-2}\, \liminf_{\varepsilon \to 0} E\big(\psi_\varepsilon(Z) Z\big) \quad \text{by (3.20)} \\
&= \sigma_{\min}^{-2}\, E\left(Z \int_{-\infty}^{Z} \mathbf{1}_{[a,b]}(u)du\right) \quad \text{by bounded convergence} \\
&= \sigma_{\min}^{-2} \int_a^b E\big(Z\mathbf{1}_{[u,+\infty)}(Z)\big)du \leqslant (b-a) \times \sigma_{\min}^{-2}\, E|Z|.
\end{aligned}$$

This implies the absolute continuity of Z, that is the existence of ρ.

Step 3: A key formula. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function with compact support, and F denote any antiderivative of f. Note that F is bounded. We have

$$\begin{aligned}
E\big(f(Z)\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}\big) &= E\big(F(Z)Z\big) \quad \text{by (3.20)} \\
&= \int_{\mathbb{R}} F(z)\, z\, \rho(z)dz \\
&\underset{(*)}{=} \int_{\mathbb{R}} f(z) \left(\int_z^\infty y\rho(y)dy\right) dz \\
&= E\left(f(Z) \frac{\int_Z^\infty y\rho(y)dy}{\rho(Z)}\right).
\end{aligned}$$

Equality (*) was obtained by integrating by parts, after observing that

$$\int_z^\infty y\rho(y)dy \longrightarrow 0 \ \text{ as } |z| \to \infty$$

(for $z \to +\infty$, this is because $Z \in L^1(\Omega)$; for $z \to -\infty$, this is because Z has mean zero). Therefore, we have shown

$$g(Z) = E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}} | Z) = \frac{\int_Z^\infty y\rho(y)dy}{\rho(Z)} \quad \text{almost surely.} \tag{3.21}$$

Step 4: The support of ρ. Since $Z \in \mathbb{D}^{1,2}$, it is known (see e.g. [12, Proposition 2.1.7]) that $\mathrm{Supp}\rho = [\alpha, \beta]$ with $-\infty \leqslant \alpha < \beta \leqslant +\infty$. Since Z has zero mean, note that $\alpha < 0$ and $\beta > 0$ necessarily. Identity (3.21) yields

$$\int_z^\infty y\rho\,(y)\,dy \geqslant \sigma_{\min}^2\, \rho\,(z) \quad \text{for almost all } z \in (\alpha, \beta). \tag{3.22}$$

For every $z \in (\alpha, \beta)$, define $\varphi(z) := \int_z^\infty y\rho(y)\,dy$. This function is differentiable almost everywhere on (α, β), and its derivative is $-z\rho(z)$. In particular, since $\varphi(\alpha) = \varphi(\beta) = 0$, we have that $\varphi(z) > 0$ for all $z \in (\alpha, \beta)$. On the other hand, when multiplied by $z \in [0, \beta)$, the inequality (3.22) gives $\frac{\varphi'(z)}{\varphi(z)} \geqslant -\frac{z}{\sigma_{\min}^2}$. Integrating this relation over the interval $[0, z]$ yields $\log\varphi(z) - \log\varphi(0) \geqslant -\frac{z^2}{2\,\sigma_{\min}^2}$, i.e., since $0 = E(Z) = E(Z_+) - E(Z_-)$ so that $E|Z| = E(Z_+) + E(Z_-) = 2E(Z_+) = 2\varphi(0)$, we have

$$\varphi(z) = \int_z^\infty y\rho(y)\,dy \geqslant \frac{1}{2}E|Z|e^{-\frac{z^2}{2\,\sigma_{\min}^2}}. \tag{3.23}$$

Similarly, when multiplied by $z \in (\alpha, 0]$, inequality (3.22) gives $\frac{\varphi'(z)}{\varphi(z)} \leqslant -\frac{z}{\sigma_{\min}^2}$. Integrating this relation over the interval $[z, 0]$ yields $\log\varphi(0) - \log\varphi(z) \leqslant \frac{z^2}{2\,\sigma_{\min}^2}$, i.e. (3.23) still holds for $z \in (\alpha, 0]$. Now, let us prove that $\beta = +\infty$. If this were not the case, by definition, we would have $\varphi(\beta) = 0$; on the other hand, by letting z tend to β in the above inequality, because φ is continuous, we would have $\varphi(\beta) \geqslant \frac{1}{2}E|Z|e^{-\frac{\beta^2}{2\sigma_{\min}^2}} > 0$, which contradicts $\beta < +\infty$. The proof of $\alpha = -\infty$ is similar. In conclusion, we have shown that $\mathrm{supp}\rho = \mathbb{R}$.

Step 5: Proof of (3.19). Let $\varphi : \mathbb{R} \to \mathbb{R}$ be still defined by $\varphi(z) = \int_z^\infty y\rho(y)dy$. On one hand, we have $\varphi'(z) = -z\rho(z)$ for almost all $z \in \mathbb{R}$. On the other hand, by (3.21), we have, for almost all $z \in \mathbb{R}$,

$$\varphi(z) = \rho(z)g(z). \tag{3.24}$$

By putting these two facts together, we get the following ordinary differential equation satisfied by φ:

$$\frac{\varphi'(z)}{\varphi(z)} = -\frac{z}{g(z)} \quad \text{for almost all } z \in \mathbb{R}.$$

Integrating this relation over the interval $[0, z]$ yields

$$\log\varphi(z) = \log\varphi(0) - \int_0^z \frac{x\,dx}{g(x)}.$$

Taking the exponential and using the fact that $\varphi(0) = \frac{1}{2}\,E|Z|$, we get

$$\varphi(z) = \frac{1}{2}\,E|Z|\exp\left(-\int_0^z \frac{x\,dx}{g(x)}\right).$$

Finally, the desired conclusion comes from (3.24). $\square$

Remark 3.2 *The "integration by parts formula" (3.20) was proved and used for the first time by Nourdin and Peccati in [11], in order to perform error bounds in the normal approximation of Z. Specifically, [11] shows, by combining Stein's method with (3.20), that*

$$\sup_{z\in\mathbb{R}} \big|P(Z \leqslant z) - P(N \leqslant z)\big| \leqslant \frac{\sqrt{\mathrm{Var}\big(E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z)\big)}}{\mathrm{Var}(Z)}, \tag{3.25}$$

where $N \sim \mathscr{N}(0, \mathrm{Var}Z)$. In reality, the inequality stated in [11] is with $\mathrm{Var}\big(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}\big)$ *instead of* $\mathrm{Var}\big(E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z)\big)$ *on the right-hand side; but the same proof allows to write this slight improvement; it was not stated or used in [11] because it did not improve the applications therein.*

Using Theorem 3.1, we can deduce the following interesting criterion for normality, which one will compare with (3.25).

Corollary 3.3 *Let $Z \in \mathbb{D}^{1,2}$; let $g(Z) = E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z)$. Then Z is Gaussian if and only if* $\mathrm{Var}(g(Z)) = 0$.

Proof: We can assume without loss of generality that Z is centered. By (3.20) (choose $f(z) = z$), we have

$$E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}) = E(Z^2) = \mathrm{Var}Z.$$

Therefore, the condition $\mathrm{Var}(g(Z)) = 0$ is equivalent to

$$g(Z) = \mathrm{Var}Z \quad \text{almost surely.}$$

Let $Z \sim \mathscr{N}(0, \sigma^2)$. Using (3.21), we immediately check that $g(Z) = \sigma^2$ almost surely. Conversely, if $g(Z) = \sigma^2$ almost surely, then Theorem 3.1 implies that Z has a density ρ given by $\rho(z) = \frac{E|Z|}{2\sigma^2} e^{-\frac{z^2}{2\sigma^2}}$ for almost all $z \in \mathbb{R}$, from which we immediately deduce that $Z \sim \mathscr{N}(0, \sigma^2)$.

□

Observe that if $Z \sim \mathscr{N}(0, \sigma^2)$, then $E|Z| = \sqrt{2/\pi}\,\sigma$, so that the formula (3.19) for ρ agrees, of course, with the usual one in this case.

Depending on the situation, $g(Z)$ may be computable or may be estimated by hand. We cite the next corollary for situations where this is the case. However, with the exception of this corollary, the remainder of this section, starting with Proposition 3.5, provides a systematic computational technique to deal with $g(Z)$.

Corollary 3.4 *Let $Z \in \mathbb{D}^{1,2}$ with zero mean and $g(Z) := E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z)$. If there exists $\sigma_{\min}, \sigma_{\max} > 0$ such that*

$$\sigma_{\min}^2 \leqslant g(Z) \leqslant \sigma_{\max}^2 \quad \textit{almost surely,}$$

then Z has a density ρ satisfying, for almost all $z \in \mathbb{R}$

$$\frac{E|Z|}{2\,\sigma_{\min}^2}\exp\left(-\frac{z^2}{2\sigma_{\max}^2}\right) \leqslant \rho(z) \leqslant \frac{E|Z|}{2\,\sigma_{\max}^2}\exp\left(-\frac{z^2}{2\sigma_{\min}^2}\right).$$

Proof: One only needs to apply Theorem 3.1. □

3.2 Computations and examples

We now show how to compute $g(Z) := E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z)$ in practice. We then provide several examples using this computation.

Proposition 3.5 *Write $DZ = \Phi_Z(X)$ with a measurable function $\Phi_Z : \mathbb{R}^{\mathfrak{H}} \to \mathfrak{H}$. We have*

$$g(Z) = \int_0^\infty e^{-u}\,\mathbf{E}\big(\langle \Phi_Z(X), \Phi_Z(e^{-u}X + \sqrt{1-e^{-2u}}X')\rangle_{\mathfrak{H}}|Z\big)du,$$

where X' stands for an independent copy of X, and is such that X and X' are defined on the product probability space $(\Omega\times\Omega', \mathscr{F}\otimes\mathscr{F}', P\times P')$. Here $\mathbf{E}$ denotes the mathematical expectation with respect to $P\times P'$.

Proof: We follow the arguments contained in Nourdin and Peccati [11, Remark 3.6]. Without loss of generality, we can assume that $\mathfrak{H} = L^2(A, \mathscr{A}, \mu)$ where $(A, \mathscr{A})$ is a measurable space and μ is a σ-finite measure without atoms. Let us consider the chaos expansion of Z, given by $Z = \sum_{m=1}^\infty I_m(f_m)$, with $f_m \in \mathfrak{H}^{\odot m}$. Therefore $-L^{-1}Z = \sum_{m=1}^\infty \frac{1}{m} I_m(f_m)$ and

$$-D_aL^{-1}Z = \sum_{m=1}^\infty I_{m-1}(f_m(\cdot, a)), \quad a \in A.$$

On the other hand, we have $D_aZ = \sum_{m=1}^\infty mI_{m-1}(f_m(\cdot, a))$. Thus

$$\begin{aligned}\int_0^\infty e^{-u}T_u(D_aZ)du &= \int_0^\infty e^{-u}\left(\sum_{m=1}^\infty me^{-(m-1)u}I_{m-1}(f_m(\cdot, a))\right)du \\ &= \sum_{m=1}^\infty I_{m-1}(f_m(\cdot, a)).\end{aligned}$$

Consequently,

$$-DL^{-1}Z = \int_0^\infty e^{-u}T_u(DZ)du.$$

By Mehler's formula (2.17), and since $DZ = \Phi_Z(X)$ by assumption, we deduce that

$$-DL^{-1}Z = \int_0^\infty e^{-u}E'\big(\Phi_Z(e^{-u}X + \sqrt{1-e^{-2u}}X')\big)du.$$

Using $E(E'(\ldots)|Z) = \mathbf{E}(\ldots|Z)$, the desired conclusion follows.

$\square$

By combining (3.19) with Proposition 3.5, we get the formula (1.12) given in the introduction, more precisely:

Corollary 3.6 *Let $Z \in \mathbb{D}^{1,2}$ be centered, and let $\Phi_Z : \mathbb{R}^{\mathfrak{H}} \to \mathfrak{H}$ be measurable and such that $DZ = \Phi_Z(X)$. Assume that condition (3.18) holds. Then Z has a density ρ given, for almost all $z \in \mathbb{R}$, by*

$$\rho(z) = \frac{E|Z|}{2\int_0^\infty e^{-u}\,\mathbf{E}\big(\langle \Phi_Z(X), \Phi_Z(e^{-u}X + \sqrt{1-e^{-2u}}X')\rangle_{\mathfrak{H}} | Z = z\big)du} \times \exp\left(-\int_0^z \frac{x\,dx}{\int_0^\infty e^{-u}\,\mathbf{E}\big(\langle \Phi_Z(X), \Phi_Z(e^{-u}X + \sqrt{1-e^{-2u}}X')\rangle_{\mathfrak{H}} | Z = x\big)du}\right).$$

Now, we give several examples of application of this corollary.

3.2.1 First example: monotone Gaussian functional, finite case.

Let $N \sim \mathscr{N}_n(0, K)$ with $K \in \mathscr{S}_n^+(\mathbb{R})$, and $f : \mathbb{R}^n \to \mathbb{R}$ be a $\mathscr{C}^1$ function having bounded derivatives. We assume, without loss of generality, that each N_i has the form $X(h_i)$, for a certain centered isonormal process X (over some Hilbert space $\mathfrak{H}$) and certain functions $h_i \in \mathfrak{H}$. Set $Z = f(N) - E(f(N))$. The chain rule (2.14) implies that $Z \in \mathbb{D}^{1,2}$ and that $DZ = \Phi_Z(N) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(N) h_i$. Therefore

$$\langle \Phi_Z(X), \Phi_Z(e^{-u}X + \sqrt{1-e^{-2u}}X')\rangle_{\mathfrak{H}} = \sum_{i,j=1}^n K_{ij} \frac{\partial f}{\partial x_i}(N) \frac{\partial f}{\partial x_j}(e^{-u}N + \sqrt{1-e^{-2u}}N').$$

(Compare with Lemma 5.3 in Chatterjee [4]). In particular, Corollary 3.6 yields the following.

Proposition 3.7 *Let $N \sim \mathscr{N}_n(0, K)$ with $K \in \mathscr{S}_n^+(\mathbb{R})$, and $f : \mathbb{R}^n \to \mathbb{R}$ be a $\mathscr{C}^1$ function with bounded derivatives. If there exist $\alpha_i, \beta_i \geqslant 0$ such that $\alpha_i \leqslant \frac{\partial f}{\partial x_i}(x) \leqslant \beta_i$ for any $i \in \{1, \ldots, n\}$ and $x \in \mathbb{R}^n$, if $K_{ij} \geqslant 0$ for any $i, j \in \{1, \ldots, n\}$ and if $\sum_{i,j=1}^n \alpha_i\alpha_j K_{ij} > 0$, then $Z = f(N) - E(f(N))$ has a density ρ satisfying, for almost all $z \in \mathbb{R}$,*

$$\frac{E|Z|}{2\sum_{i,j=1}^n \beta_i\beta_j K_{ij}} \exp\left(-\frac{z^2}{2\sum_{i,j=1}^n \alpha_i\alpha_j K_{ij}}\right) \leqslant \rho(z) \leqslant \frac{E|Z|}{2\sum_{i,j=1}^n \alpha_i\alpha_j K_{ij}} \exp\left(-\frac{z^2}{2\sum_{i,j=1}^n \beta_i\beta_j K_{ij}}\right).$$

3.2.2 Second example: proof of Proposition 1.1.

Let $N \sim \mathscr{N}_n(0, K)$ with $K \in \mathscr{S}_n^+(\mathbb{R})$. Once again, we assume that each N_i has the form $X(h_i)$, for a certain centered isonormal process X (over some Hilbert space $\mathfrak{H}$) and certain functions $h_i \in \mathfrak{H}$. Let $Z = \max N_i - E(\max N_i)$, and set

$$I_u = \mathrm{argmax}_{1 \leqslant i \leqslant n}(e^{-u} X(h_i) + \sqrt{1 - e^{-2u}} X'(h_i)) \quad \text{for } u \geqslant 0.$$

Lemma 3.8 *For any $u \geqslant 0$, I_u is a well-defined random element of $\{1, \ldots, n\}$. Moreover, $Z \in \mathbb{D}^{1,2}$ and we have $DZ = \Phi_Z(N) = h_{I_0}$.*

Proof: Fix $u \geqslant 0$. Since, for any $i \neq j$, we have

$$\begin{aligned} & P\big(e^{-u} X(h_i) + \sqrt{1 - e^{-2u}} X'(h_i) = e^{-u} X(h_j) + \sqrt{1 - e^{-2u}} X'(h_j)\big) \\ & = P\big(X(h_i) = X(h_j)\big) = 0, \end{aligned}$$

the random variable I_u is a well-defined element of $\{1, \ldots, n\}$. Now, if Δ_i denotes the set $\{x \in \mathbb{R}^n : x_j \leqslant x_i \text{ for all } j\}$, observe that $\frac{\partial}{\partial x_i} \max = \mathbf{1}_{\Delta_i}$ almost everywhere. The desired conclusion follows from the Lipshitz version of the chain rule (2.14), and the following Lipshitz property of the max function, which is easily proved by induction on $n \geqslant 1$:

$$\big|\max(y_1, \ldots, y_n) - \max(x_1, \ldots, x_n)\big| \leqslant \sum_{i=1}^{n} |y_i - x_i| \quad \text{for any } x, y \in \mathbb{R}^n. \tag{3.26}$$

$\square$

In particular, we deduce from Lemma 3.8 that

$$\langle \Phi_Z(X), \Phi_Z(e^{-u} X + \sqrt{1 - e^{-2u}} X') \rangle_{\mathfrak{H}} = K_{I_0, I_u}. \tag{3.27}$$

By combining this fact with Corollary 3.6, we get Proposition 1.1, which we restate.

Proposition 3.9 *Let $N \sim \mathscr{N}_n(0, K)$ with $K \in \mathscr{S}_n^+(\mathbb{R})$. If there exists $\sigma_{\min}, \sigma_{\max} > 0$ such that $\sigma_{\min}^2 \leqslant K_{ij} \leqslant \sigma_{\max}^2$ for any $i, j \in \{1, \ldots, n\}$, then $Z = \max N_i - E(\max N_i)$ has a density ρ satisfying (1.5) for almost all $z \in \mathbb{R}$.*

3.2.3 Third example: monotone Gaussian functional, continuous case.

Assume that $X = (X_t, t \in [0, T])$ is a centered Gaussian process with continuous paths, and that $f : \mathbb{R} \to \mathbb{R}$ is $\mathscr{C}^1$ with a bounded derivative. Consider $Z = \int_0^T f(X_v) dv - E\left(\int_0^T f(X_v) dv\right)$. Then $Z \in \mathbb{D}^{1,2}$ and we have $DZ = \Phi_Z(X) = \int_0^T f'(X_v) \mathbf{1}_{[0,v]} dv$. Therefore

$$\begin{aligned} & \langle \Phi_Z(X), \Phi_Z(e^{-u} X + \sqrt{1 - e^{-2u}} X') \rangle_{\mathfrak{H}} \\ & = \iint_{[0,T]^2} f'(X_v) f'(e^{-u} X_w + \sqrt{1 - e^{-2u}} X'_w) E(X_v X_w) dv dw. \end{aligned}$$

Using Corollary 3.6, we get the following.

Proposition 3.10 *Assume that* $X = (X_t,\, t \in [0,T])$ *is a centered Gaussian process with continuous paths, and that* $f : \mathbb{R} \to \mathbb{R}$ *is* $\mathscr{C}^1$. *If there exists* $\alpha, \beta, \sigma_{\min}, \sigma_{\max} > 0$ *such that* $\alpha \leqslant f'(x) \leqslant \beta$ *for all* $x \in \mathbb{R}$ *and* $\sigma^2_{\min} \leqslant E(X_v X_w) \leqslant \sigma^2_{\max}$ *for all* $v, w \in [0,T]$, *then* $Z = \int_0^T f(X_v)dv - E\left(\int_0^T f(X_v)dv\right)$ *has a density* ρ *satisfying, for almost all* $z \in \mathbb{R}$,

$$\frac{E|Z|}{2\beta^2\,\sigma^2_{\max}\,T^2}\, e^{-\frac{z^2}{2\alpha^2\,\sigma^2_{\min}T^2}} \leqslant \rho(z) \leqslant \frac{E|Z|}{2\alpha^2\,\sigma^2_{\min}\,T^2}\, e^{-\frac{z^2}{2\beta^2\,\sigma^2_{\max}T^2}}.$$

3.2.4 Fourth example: supremum of a Gaussian process

Fix $a < b$, and assume that $X = (X_t,\, t \in [a,b])$ is a centered Gaussian process with continuous paths and such that $E|X_t - X_s|^2 \neq 0$ for all $s \neq t$. Set $Z = \sup_{[a,b]} X - E(\sup_{[a,b]} X)$, and let τ_u be the (unique) random point where $e^{-u}X + \sqrt{1-e^{-2u}}X'$ attains its maximum on $[a,b]$. Note that τ_u is well-defined, see e.g. Lemma 2.6 in [7]. Moreover, we have that $Z \in \mathbb{D}^{1,2}$, see Proposition 2.1.10 in [12], and $DZ = \Phi_Z(X) = \mathbf{1}_{[0,\tau_0]}$, see Lemma 3.1 in [5]. Therefore

$$\langle \Phi_Z(X), \Phi_Z(e^{-u}X + \sqrt{1-e^{-2u}}X')\rangle_{\mathfrak{H}} = R(\tau_0, \tau_u)$$

where $R(s,t) = E(X_s X_t)$ is the covariance function of X. Using Corollary 3.6, the following obtains.

Proposition 3.11 *Let* $X = (X_t,\, t \in [a,b])$ *be a centered Gaussian process with continuous paths, and such that* $E|X_t - X_s|^2 \neq 0$ *for all* $s \neq t$. *Assume that, for some real* $\sigma_{\min}, \sigma_{\max} > 0$, *we have* $\sigma^2_{\min} \leqslant E(X_s X_t) \leqslant \sigma^2_{\max}$ *for any* $s, t \in [a,b]$. *Then,* $Z = \sup_{[a,b]} X - E(\sup_{[a,b]} X)$ *has a density* ρ *satisfying, for almost all* $z \in \mathbb{R}$,

$$\frac{E|Z|}{2\sigma^2_{\max}}\, e^{-\frac{z^2}{2\,\sigma^2_{\min}}} \leqslant \rho(z) \leqslant \frac{E|Z|}{2\sigma^2_{\min}}\, e^{-\frac{z^2}{2\,\sigma^2_{\max}}}.$$

To the best of our knowledge, Proposition 3.11, as well as Proposition 3.9, contain the first bounds ever established for the *density* of the supremum of a general Gaussian process. When integrated over z, the upper bound above improves the classical concentration inequalities (1.2), (1.3), (1.4) on the tail of Z, see e.g. the upper bound in (1.9); the lower bound for the left-hand tail of Z which one obtains by integration, appears to be entirely new. When applied to the case of fractional Brownian motion, we get the following.

Corollary 3.12 *Let* $b > a > 0$, *and* $B = (B_t,\, t \geqslant 0)$ *be a fractional Brownian motion with Hurst index* $H \in [1/2, 1)$. *Then the random variable* $Z = \sup_{[a,b]} B - E\big(\sup_{[a,b]} B\big)$ *has a density* ρ *satisfying (1.6) for almost all* $z \in \mathbb{R}$.

Proof: For any choice of the Hurst parameter $H \in (1/2, 1)$, the Gaussian space generated by B can be identified with an isonormal Gaussian process of the type $X = \{X(h) :$

$h \in \mathfrak{H}\}$, where the real and separable Hilbert space $\mathfrak{H}$ is defined as follows: (i) denote by $\mathscr{E}$ the set of all $\mathbb{R}$-valued step functions on $\mathbb{R}_+$, (ii) define $\mathfrak{H}$ as the Hilbert space obtained by closing $\mathscr{E}$ with respect to the scalar product

$$\left\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \right\rangle_{\mathfrak{H}} = E(B_t B_s) = \frac{1}{2}\big(t^{2H} + s^{2H} - |t-s|^{2H}\big).$$

In particular, with such a notation, one has that $B_t = X(\mathbf{1}_{[0,t]})$. The reader is referred e.g. to [12] for more details on fractional Brownian motion.

Now, the desired conclusion is a direct application of Proposition 3.11 since, for all $a \leqslant s < t \leqslant b$,

$$E(B_s B_t) \leqslant \sqrt{E(B_s^2)}\sqrt{E(B_t^2)} = (st)^H \leqslant b^{2H}$$

and

$$\begin{aligned} E(B_s B_t) &= \frac{1}{2}\big(t^{2H} + s^{2H} - (t-s)^{2H}\big) = H(2H-1)\iint_{[0,s]\times[0,t]} |v-u|^{2H-2} du dv \\ &\geqslant H(2H-1)\iint_{[0,a]\times[0,a]} |v-u|^{2H-2} du dv = E(B_a^2) = a^{2H}. \end{aligned}$$

□

4 Concentration inequalities

Now, we investigate what can be said when $g(Z) = E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z)$ just admits a lower (resp. upper) bound. Results under such hypotheses are more difficult to obtain than in the previous section, since there we could use bounds on $g(Z)$ in both directions to good effect; this is apparent, for instance, in the appearance of both the lower and upper bounding values $\sigma_{\min}$ and $\sigma_{\max}$ in each of the two bound in (1.5), or more generally in Corollary 3.4. However, given our previous work, tails bounds can be readily obtained: most of the analysis of the role of $g(Z)$ in tail estimates is already contained in the proof of Theorem 3.1.

4.1 Upper bounds

Our first result allows comparisons both to the Gaussian and exponential tails.

Theorem 4.1 *Let $Z \in \mathbb{D}^{1,2}$ with zero mean, $g(Z) = E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z)$, and fix $\alpha \geqslant 0$ and $\beta > 0$. Assume that*

(i) $g(Z) \leqslant \alpha Z + \beta$ almost surely;

(ii) Z has a density ρ.

Then, for all $z > 0$, we have

$$P(Z \geqslant z) \leqslant \exp\left(-\frac{z^2}{2\alpha z + 2\beta}\right).$$

Proof: We follow the same line of reasoning as in [3, Theorem 1.5]. For any $A > 0$, define $m_A : [0, +\infty) \to \mathbb{R}$ by $m_A(\theta) = E\big(e^{\theta Z}\mathbf{1}_{\{Z \leqslant A\}}\big)$. By Lebesgue differentiation theorem, we have

$$m_A'(\theta) = E(Ze^{\theta Z}\mathbf{1}_{\{Z\leqslant A\}}) \quad \text{for all } \theta \geqslant 0.$$

Therefore, we can write

$$\begin{aligned} m_A'(\theta) &= \int_{-\infty}^{A} z\, e^{\theta z}\, \rho(z) dz \\ &= -e^{\theta A}\int_A^\infty y\rho(y)dy + \theta\int_{-\infty}^{A} e^{\theta z}\left(\int_z^\infty y\rho(y)dy\right)dz \quad \text{by integration by parts} \\ &\leqslant \theta\int_{-\infty}^{A} e^{\theta z}\left(\int_z^\infty y\rho(y)dy\right)dz \quad \text{since } \textstyle\int_A^\infty y\rho(y)dy \geqslant 0 \\ &= \theta E\big(g(Z)\, e^{\theta Z}\, \mathbf{1}_{\{Z\leqslant A\}}\big), \end{aligned}$$

where the last line follows from identity (3.21). Due to the assumption (i), we get

$$m_A'(\theta) \leqslant \theta\,\alpha\, m_A'(\theta) + \theta\,\beta\, m_A(\theta),$$

that is, for any $\theta \in (0, 1/\alpha)$:

$$\frac{m_A'(\theta)}{m_A(\theta)} \leqslant \frac{\theta\beta}{1-\theta\alpha}.$$

By integration and since $m_A(0) = P(Z \leqslant A) \leqslant 1$, this gives, for any $\theta \in (0, 1/\alpha)$:

$$m_A(\theta) \leqslant \exp\left(\int_0^\theta \frac{\beta u}{1-\alpha u}du\right) \leqslant \exp\left(\frac{\beta\theta^2}{2(1-\theta\alpha)}\right).$$

Using Fatou's inequality (as $A \to \infty$) in the previous relation implies

$$E\big(e^{\theta Z}\big) \leqslant \exp\left(\frac{\beta\theta^2}{2(1-\theta\alpha)}\right)$$

for all $\theta \in (0, 1/\alpha)$. Therefore, for all $\theta \in (0, 1/\alpha)$, we have

$$P(Z \geqslant z) = P(e^{\theta Z} \geqslant e^{\theta z}) \leqslant e^{-\theta z}E\big(e^{\theta Z}\big) \leqslant \exp\left(\frac{\beta\theta^2}{2(1-\theta\alpha)} - \theta z\right).$$

Choosing $\theta = \frac{z}{\alpha z+\beta} \in (0, 1/\alpha)$ gives the desired result.

$\square$

Let us give an example of application of Theorem 4.1. Assume that $B = (B_t,\, t \geqslant 0)$ is a fractional Brownian motion with Hurst index $H \in (0,1)$. For any choice of the parameter H, as already mentioned in the proof of Corollary 3.12, the Gaussian space generated by B can be identified with an isonormal Gaussian process of the type $X = \{X(h) : h \in \mathfrak{H}\}$, where the real and separable Hilbert space $\mathfrak{H}$ is defined as follows: (i) denote by $\mathscr{E}$ the set of all $\mathbb{R}$-valued step functions on $\mathbb{R}_+$, (ii) define $\mathfrak{H}$ as the Hilbert space obtained by closing $\mathscr{E}$ with respect to the scalar product

$$\left\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \right\rangle_{\mathfrak{H}} = E(B_t B_s) = \frac{1}{2}\big(t^{2H} + s^{2H} - |t-s|^{2H}\big).$$

In particular, with such a notation one has that $B_t = X(\mathbf{1}_{[0,t]})$. Now, let

$$Z = Z_T := \int_0^T B_u^2 du - \frac{T^{2H+1}}{2H+1}.$$

By the scaling property of fractional Brownian motion, we see first that Z_T has the same distribution as $T^{2H+1} Z_1$. Thus we choose $T = 1$ without loss of generality; we denote $Z = Z_1$. Now observe that $Z \in \mathbb{D}^{1,2}$ lives in the second Wiener chaos of B. In particular, we have $-L^{-1}Z = \frac{1}{2}Z$. Moreover $DZ = 2\int_0^1 B_u \, \mathbf{1}_{[0,u]} du$, so that

$$\begin{aligned}
\langle DZ, -DL^{-1}Z \rangle_{\mathfrak{H}} &= \frac{1}{2}\|DZ\|_{\mathfrak{H}}^2 = 2\int_{[0,1]^2} B_u B_v \, E(B_u B_v) du dv \\
&\leqslant 2\int_{[0,1]^2} |B_u|\,|B_v|\,\big|E(B_u B_v)\big| du dv \\
&\leqslant 2\int_{[0,1]^2} |B_u|\,|B_v|\, u^H v^H du dv = 2\left(\int_0^1 |B_u| u^H du\right)^2 \\
&\leqslant 2\int_0^1 B_u^2 du \times \int_0^1 u^{2H} du = \frac{1}{H+1/2}\int_0^1 B_u^2 du \\
&= \frac{1}{H+1/2}\left(Z + \frac{1}{2H+1}\right).
\end{aligned}$$

Since it is easily shown that Z has a density, Theorem 4.1 implies the desired conclusion in Proposition 1.3, or with $c_H = H + 1/2$,

$$P\left(Z_1 \geqslant z\right) \leqslant \exp\left(-\frac{z^2 c_H^2}{2c_H z + 1}\right).$$

By scaling, this shows that the tail of Z_T / T^{2H+1} behaves asymptotically like that of an exponential random variable with mean $\nu = (H/2 + 1/4)^{-1}$.

For the moment, it is not possible to use our tools to investigate a lower bound on this tail, see the forthcoming Section 4.2. We have also investigated the possibility of using such tools as the formula (1.10), or the density lower bounds found in [13], thinking that a specific second-chaos situation might be tractable despite the reliance on the divergence operator, but these tools seem even less appropriate. However, in this particular instance, we can perform a calculation by hand, as follows. By Jensen's inequality, with $\mu = (2H+1)^{-1}$, we have that $Z + \mu = Z_1 + \mu \geqslant \left(\int_0^1 B_u du\right)^2$. Thus

$$P\big(Z_1 \geqslant z\big) \geqslant P\left(\left(\int_0^1 B_u du\right)^2 \geqslant z + \mu\right) = P\left(\left|\int_0^1 B_u du\right| \geqslant \sqrt{z+\mu}\right).$$

Here of course, the random variable $N = \int_0^1 B_u du$ is centered Gaussian, and its variance can be calculated by hand:

$$\begin{aligned}\sigma^2 := E\left(N^2\right) &= \iint_{[0,1]^2} E\left(B_u B_v\right) du dv \\ &= \int_0^1 dv \int_0^v du \left(u^{2H} + v^{2H} - (v-u)^{2H}\right) = \frac{1}{2H+2}.\end{aligned}$$

Therefore, by the standard lower bound on the tail of a Gaussian r.v., that is $\int_z^\infty e^{-y^2/2} dy \geqslant \frac{z}{1+z^2} e^{-z^2/2}$ for all $z > 0$, we get

$$\begin{aligned}P\left(Z_1 \geqslant z\right) &\geqslant \frac{\sigma\sqrt{z+\mu}}{\sigma^2 + z + \mu} \exp\left(-\frac{z+\mu}{2\sigma^2}\right) \\ &\underset{z\to\infty}{\sim} \frac{1}{\sqrt{z}\sqrt{2H+2}} \exp\left(-\frac{H+1}{2H+1}\right) \exp\left(-\left(H+1\right)z\right). \qquad (4.28)\end{aligned}$$

Abusively ignoring the factor $z^{-1/2}$ in this lower bound, we can summarize our results by saying that Z_T / T^{2H+1} has a tail that is bounded above and below by exponential tails with respective means $(H/2 + 1/4)^{-1}$ and $(H+1)^{-1}$.

As another example, let us explain how Theorem 4.1 allows to easily recover both the Borell-Sudakov-type inequalities (1.2) and (1.3), for Z defined as the centered supremum of a Gaussian vector in (1.1). We can assume, without loss of generality, that each N_i has the form $X(h_i)$, for a certain centered isonormal process X (over some Hilbert space $\mathfrak{H}$) and certain functions $h_i \in \mathfrak{H}$. Condition (ii) of Theorem 4.1 is easily satisfied while for condition (i), we have, by combining (3.27) with Proposition 3.5:

$$\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}} = \int_0^\infty e^{-u} K_{I_0, I_u} du \leqslant \max_{1\leqslant i,j\leqslant n} K_{ij} = \sigma^2_{\max} \qquad (4.29)$$

so that

$$g(Z) \leqslant \sigma^2_{\max} \quad \text{almost surely.}$$

In other words, condition (i) is satisfied with $\alpha = 0$ and $\beta = \sigma^2_{\max}$. Therefore $P(Z \geqslant z) \leqslant \exp\left(-\frac{z^2}{2\,\sigma^2_{\max}}\right)$, for all $z > 0$, and (1.2) is shown. The proof of (1.3) follows the same lines, by considering $-Z$ instead of Z.

4.2 Lower bounds

We now investigate a lower bound analogue of Theorem 4.1. Recall we still use the notation $g(z) = E(\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}} | Z = z)$.

Theorem 4.2 *Let $Z \in \mathbb{D}^{1,2}$ with zero mean, and fix $\sigma_{\min}, \alpha > 0$ and $\beta > 1$. Assume that*

(i) $g(Z) \geqslant \sigma_{\min}^2$ almost surely.

The existence of the density ρ of Z is thus ensured by Theorem 3.1. Also assume that

(ii) the function $h(x) := x^{1+\beta}\rho(x)$ is decreasing on $[\alpha, +\infty)$.

Then, for all $z \geqslant \alpha$, we have

$$P(Z \geqslant z) \geqslant \frac{1}{2}\left(1 - \frac{1}{\beta}\right) E|Z| \, \frac{1}{z} \exp\left(-\frac{z^2}{2\,\sigma_{\min}^2}\right).$$

Alternately, instead of (ii), assume that there exists $0 < \alpha < 2$ such that

(ii)' $\limsup_{z\to\infty} z^{-\alpha} \log g(z) < \infty$.

Then, for any $\varepsilon > 0$, there exist $K, z_0 > 0$ such that, for all $z > z_0$,

$$P(Z \geqslant z) \geqslant K \, \exp\left(-\frac{z^2}{(2-\varepsilon)\,\sigma_{\min}^2}\right).$$

Proof: First, let us relate the function $\varphi(z) = \int_z^\infty y\rho(y)dy$ to the tail of Z. By integration by parts, we get

$$\varphi(z) = z\, P(Z \geqslant z) + \int_z^\infty P(Z \geqslant y)dy. \tag{4.30}$$

If we assume (ii), since h is decreasing, for any $y > z \geqslant \alpha$ we have $\frac{y\rho(y)}{z\rho(z)} \leqslant \left(\frac{z}{y}\right)^\beta$. Then we have, for any $z \geqslant \alpha$:

$$P(Z \geqslant z) = z\rho(z) \int_z^\infty \frac{1}{y} \frac{y\rho(y)}{z\rho(z)} dy \leqslant z\rho(z)\, z^\beta \int_z^\infty \frac{dy}{y^{1+\beta}} = \frac{z\rho(z)}{\beta}.$$

By putting that inequality into (4.30), we get

$$\varphi(z) \leqslant z\, P(Z \geqslant z) + \frac{1}{\beta}\int_z^\infty y\rho(y)dy = z\, P(Z \geqslant z) + \frac{1}{\beta}\varphi(z)$$

so that $P(Z \geqslant z) \geqslant \left(1 - \frac{1}{\beta}\right)\frac{\varphi(z)}{z}$. Combined with (3.23), this gives the desired conclusion.

Now assume $(ii)'$ instead. Here the proof needs to be modified. From the key result of Theorem 3.1 and condition (i), we have

$$\rho(z) \geqslant \frac{E|Z|}{2\,g(z)} \exp\left(-\frac{z^2}{2\,\sigma_{\min}^2}\right).$$

Let $\Psi(z)$ denote the unnormalized Gaussian tail $\int_z^\infty \exp\left(-\frac{y^2}{2\,\sigma_{\min}^2}\right) dy$. We can write, using the Schwarz inequality,

$$\begin{aligned}\Psi^2(z) &= \left(\int_z^\infty \exp\left(-\frac{y^2}{2\,\sigma_{\min}^2}\right) \sqrt{g(y)}\,\frac{1}{\sqrt{g(y)}}\,dy\right)^2 \\ &\leqslant \int_z^\infty \exp\left(-\frac{y^2}{2\,\sigma_{\min}^2}\right) g(y)\,dy \times \int_z^\infty \exp\left(-\frac{y^2}{2\,\sigma_{\min}^2}\right) \frac{1}{g(y)}\,dy\end{aligned}$$

so that

$$\begin{aligned}P(Z \geqslant z) &= \int_z^\infty \rho(y)\,dy \\ &\geqslant \frac{E\,|Z|}{2} \int_z^\infty e^{-y^2/\left(2\sigma_{\min}^2\right)}\,\frac{1}{g(y)}\,dy \\ &\geqslant \frac{E\,|Z|}{2}\,\frac{\Psi^2(z)}{\int_z^\infty e^{-y^2/\left(2\sigma_{\min}^2\right)} g(y)\,dy}.\end{aligned}$$

Using the classical inequality $\int_z^\infty e^{-y^2/2} dy \geqslant \frac{z}{1+z^2} e^{-z^2/2}$, we get

$$P(Z \geqslant z) \geqslant \frac{E|Z|}{2}\,\frac{\sigma_{\min}^4 z^2}{\left(\sigma_{\min}^2 + z^2\right)^2}\,\frac{\exp\left(-\frac{z^2}{\sigma_{\min}^2}\right)}{\int_z^\infty \exp\left(-\frac{y^2}{2\sigma_{\min}^2}\right) g(y)dy}. \tag{4.31}$$

Under condition $(ii)'$, we have that there exists $c > 0$ such that, for y large enough, $g(y) \leqslant e^{cy^\alpha}$ with $0 < \alpha < 2$. We leave it to the reader to check that the conclusion now follows by an elementary calculation from (4.31). □

Remark 4.3 *1. Inequality (4.31) itself may be of independent interest, when the growth of g can be controlled, but not as efficiently as in $(ii)'$.*

2. Condition (ii) implies that Z has a moment of order greater than β. Therefore it can be considered as a technical regularity and integrability condition. Condition $(ii)'$ may be easier to satisfy in cases where a good handle on g exists. Yet the use of the Schwarz inequality in the above proof means that conditions $(ii)'$ is presumably stronger than it needs to be.

3. *In general, one can see that deriving lower bounds on tails of random variables with little upper bound control is a difficult task, deserving of further study.*

Acknowledgment: We are grateful to Paul Malliavin, David Nualart, and Giovanni Peccati, for helpful comments.

References

[1] Borell, Ch. (1978). Tail probabilities in Gauss space. In *Vector Space Measures and Applications, Dublin, 1977. Lecture Notes in Math.* **644**, 71-82. Springer-Verlag.

[2] Borell, Ch. (2006). On a certain exponential inequality for Gaussian processes. *Extremes* **9**, 169-176.

[3] Chatterjee, S. (2008). Stein's method for concentration inequalities. *Probab. Theory Rel. Fields*, to appear.

[4] Chatterjee, S. (2008). Fluctuation of eigenvalues and second order Poincaré inequalities. *Probab. Theory Rel. Fields*, to appear.

[5] Decreusefond, L.; Nualart, D. (2008). Hitting times for Gaussian processes. *Ann. Probab.* **36** (1), 319-330.

[6] Houdré, C.; Privault, N. (2002). Concentration and deviation inequalities in infinite dimensions via covariance representations. *Bernoulli* **8** (6), 697-720.

[7] Kim, J.; Pollard, D. (1990). Cube root asymptotics. *Ann. Probab.* **18**, 191-219.

[8] Kohatsu-Higa, A. (2003). Lower bounds for densities of uniformly elliptic random variables on Wiener space. *Probab. Theory Relat. Fields* **126**, 421-457.

[9] Kusuoka, S.; Stroock, D. (1987). Applications of the Malliavin Calculus, Part III. *J. Fac. Sci. Univ. Tokyo Sect IA Math.* **34**, 391–442.

[10] Ledoux, M.; Talagrand, M. (1991). *Probability on Banach Spaces*. Springer.

[11] Nourdin, I.; Peccati, G. (2008). Stein's method on Wiener chaos. *Probab. Theory Rel. Fields*, to appear.

[12] Nualart, D. (2006). *The Malliavin calculus and related topics*. Springer Verlag, Berlin, Second edition.

[13] Nualart, E. (2004). Exponential divergence estimates and heat kernel tail. *C. R. Math. Acad. Sci. Paris* **338** (1), 77-80.

[14] Pisier, G. (1986). Probabilistic methods in the geometry of Banach spaces. In *G. Letta and M. Pratelli (eds), Probability and Analysis, Lecture Notes in Math.* **1206**, 167-241. Springer-Verlag.

[15] Viens, F.; Vizcarra, A. (2007). Supremum Concentration Inequality and Modulus of Continuity for Sub-nth Chaos Processes. *J. Funct. Anal.* **248** 1-26.

[16] Vitale, R. (1996). The Wills functional and Gaussian processes. *Ann. Probab.* **24** (4), 2172-2178.